DOCUMENT RESUME

ED 329 877 CG 023 248

AUTHOR Shoenfelt, Elizabeth L.; Mendel, R. Max
TITLE Gender Bias in the Evaluation of Male and Female Police Officer Performance.
PUB DATE 21 Mar 91
NOTE 8p.; Paper presented at the Annual Meeting of the Southeastern Psychological Association (37th, New Orleans, LA, March 20-23, 1991).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Evaluation Criteria; Expectation; Performance Factors; *Personnel Evaluation; *Police; *Sex Bias; Sex Discrimination

ABSTRACT

This paper reports on a study undertaken to evaluate the likelihood of anti-female bias in performance appraisals. Based on investigative reports, depositions, and eye-witness accounts in the subsequent litigation, four versions of a written scenario were prepared of an actual arrest sequence in which a male training officer was shot, and the female trainee-partner was fired for cowardice. The names and pronouns used to describe the two officers were altered. One of the four scenarios was randomly selected and sent to the police chief in each of the 226 U.S. cities with populations greater than 80 thousand. The respondents were asked to complete a behaviorally-anchored performance rating form by selecting one of seven possible administrative actions ranging from termination to a meritorious performance award recommendation. The data revealed that police departments did permit gender to influence their assessment. It was found that for precisely the same actions, police departments were significantly more likely to terminate female trainees than male trainees. Given the inconsistent pattern of findings in the literature, the demonstration of gender bias in this study suggets three potential moderators worthy of further investigation. Gender bias may be more likely: (1) in situations involving strongly gender stereotyped performance settings (e.g., dangerous situations involving physically demanding actions); (2) where team performance forces the rater to apportion responsibility for outcomes among team members; and (3) where performance judgments are directly linked to specific administrative actions. (LLL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED329877

GENDER BIAS IN THE EVALUATION OF MALE AND FEMALE POLICE OFFICER PERFORMANCE

R. Max Mendel
&
Elizabeth Shoenfelt

Department of Psychology
Western Kentucky University
Bowling Green, Kentucky 42101

CG023248

Presented at The Annual Convention of the
Southeastern Psychological Association
New Orleans, Louisiana
March 21, 1991

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Elizabeth Shoenfelt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

GENDER BIAS IN THE EVALUATION OF MALE AND FEMALE POLICE OFFICER PERFORMANCE

Following an arrest attempt in which a male training officer was shot, the female trainee-partner was fired for "cowardice." The authors were contacted to assist in the evaluation of the fired female officer's Title VII claim that she was the victim of a biased appraisal process. A review of the literature on anti-female bias in performance appraisals revealed conflicting results, suggesting a number of possible moderators of the effect, but providing little empirical basis for a definitive opinion in the instant case. Accordingly, the present study was undertaken to evaluate the likelihood of anti-female bias in the present set of factual circumstances. Specifically, is a female compared to a male trainee likely to be more harshly evaluated in the present circumstances?

Based on investigative reports, depositions, and eye-witness accounts in the above litigation, four versions of a written scenario were prepared which chronicled the actual arrest sequence involving the two police officers and the suspect. The four scenarios were identical except that the names and pronouns used to describe the two officers were altered to produce the four possible gender pairings.

A behaviorally-anchored performance rating form was developed which required the respondent to evaluate the performance of each of the two officers five times by selecting one of seven possible administrative actions ranging from termination to a meritorious performance award recommendation. The five evaluations consisted of an evaluation of each of four phases of the arrest sequence as well as an overall evaluation.

One of the four scenarios was randomly selected and then sent to the police chief in each of the 226 U.S. cities with populations greater than eighty thousand. Instructions indicated that the instrument should be completed either by the chief or some other senior officer familiar with proper police procedures and experienced in the evaluation of police officer performance. Respondents were told only that our purpose was to

explore "human decision processes in the performance evaluation context." No indication of the litigious underpinnings of the study was provided.

One hundred fifty-seven (70%) scenario ratings were returned. The results of a 2 (field training officer gender) by 2 (trainee gender) ANOVA of the training officer (FTO) overall performance ratings is shown in Table 1. A significant interaction was detected.

Table 1. Anova Summary Table for FTO Overall Performance Rating

Source of Variation	DF	Mean Square	F	Signif of F
FTO	1	.370	.258	.612
TRAINEE	1	4.811	3.349	.069
FTO x TRAINEE	1	9.146	6.365	.013
Residual	153	1.437		

The analysis reveals that while neither the gender of the FTO nor the gender of the trainee alone had a significant effect on the FTO rating, the particular gender mix of the team, the FTO by Trainee interaction, did significantly impact the FTO ratings. The means for each gender combination are shown in Table 2.

Table 2. Mean FTO Performance Rating by Treatment Condition

	Male FTO	Female FTO
Male Trainee	2.51	2.92
Female Trainee	3.38	2.81

A post-hoc analysis revealed what is apparent from inspection of the cell means. The performance of the FTO is rated significantly higher when a male FTO is paired with a female trainee than in any other gender combination.

A parallel two-way ANOVA was performed on the trainee overall performance rating. Tables 3 and 4 reveal that the trainee overall performance rating was not effected by the gender mix of the arresting team. The means displayed in Table 4 are similar across the four conditions.

Table 3. ANOVA Summary Table for Trainee Overall Performance Rating

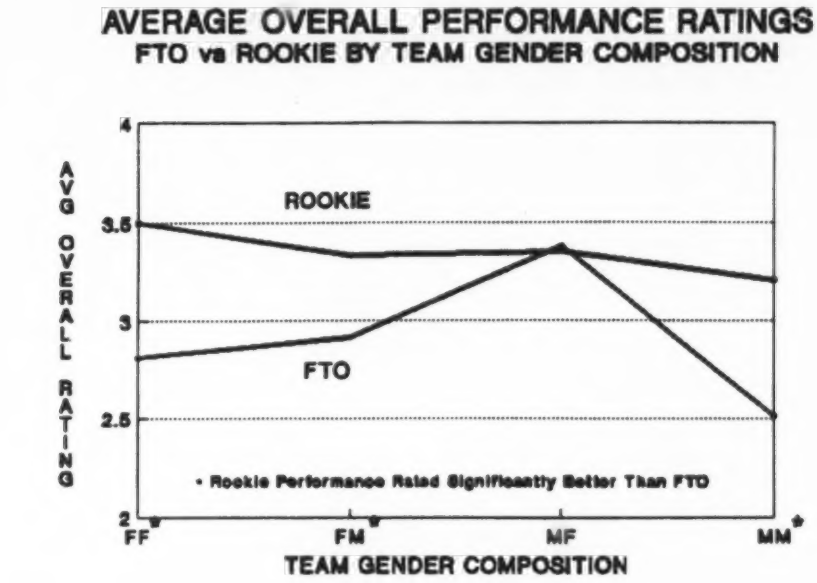
Source of Variation	DF	Mean Square	F	Signif of F
FTO	1	.735	.468	.495
TRAINEE	1	.965	.615	.434
FTO x TRAINEE	1	.003	.002	.963
Residual	153	1.569		

Table 4. Mean Trainee Performance Rating by Treatment Condition

	Male FTO	Female FTO
Male Trainee	3.21	3.33
Female Trainee	3.35	3.50

Together the results displayed in Tables 1 through Table 4 indicate the gender composition of the arresting team does indeed impact the relative performance evaluation of team members. Though the trainee's evaluation is not directly effected, the FTO performance evaluation is elevated when the FTO is male and the trainee is female. In this condition alone, the male FTO's performance is rated significantly higher than in any other gender combination. This effect is graphically displayed in Figure 1.

Figure 1. Plot of FTO and Trainee Mean Performance Rating by Team Gender Composition



Light is also shed on the related question of whether the female trainee is relatively more harshly evaluated. Recall that we did not find that the gender of the trainee directly affects the trainee evaluation. However, four paired-samples t-tests confirmed that the trainee's performance is regarded as significantly <u>better</u> than the FTO's performance in three of the four conditions ($p<.01$). The one gender combination where this relationship does not occur is when a female trainee is paired with a male FTO. In this team, the very same male FTO behaviors are evaluated more favorably than in the other three conditions. This finding was particularly relevant in the present context because the one condition where the trainee's performance suffers by contrast with the FTO's performance is precisely that gender combination that was involved in the arrest incident that led to the female trainee's termination.

An additional analysis focused on the decision to terminate the trainee, and whether this decision is made without regard to the trainee's gender. A contingency analysis was performed on the dichotomized survey ratings of the trainee's overall performance. Consistent with the rating scale anchors, trainee performance ratings of "2" or greater were coded as a "retain" decision, ratings of "1" were coded as a termination decision. We then examined the frequency with which female trainees were terminated compared to male trainees. Table 5 displays both the obtained frequency and in parentheses, the frequency one would expect given a gender neutral evaluation process.

While only 7% of the respondents indicated the trainee should be terminated, a disproportionate number of the terminations occurred for female trainees. Of 11 trainee termination decisions, nine of these were rendered for female trainees. Only two termination recommendations were returned for male trainees.

Table 5. Contingency Analyses of Termination/Retention Decisions by Gender of Trainee

	FIRED	RETAINED
FEMALE	9	73
MALE	2	73

To examine whether the obtained departure from the expected frequencies was sufficiently large to confirm the hypothesis that female trainees are more frequently terminated than male trainees, a chi-square test of statistical significance was performed. The chi-square test was significant ($X^2=4.15$, $p<.05$). Police departments were more inclined to discharge female than male trainees for precisely the same behaviors.

Taken together, the data from this study raise serious questions about the influence of sex bias in police officer performance evaluations. The

data show that police departments, perhaps quite unconsciously, do permit gender to influence their assessment of the effectiveness of an officer's performance. Precisely the same behaviors in the arrest scenario were evaluated differently depending on the gender combination of the team. The data show that the performance of the trainee is regarded as superior to that of the FTO except when the FTO is male and the trainee is female. In this latter case, the female trainee is evaluated less favorably than the male FTO.

More disturbing is the finding that, for precisely the same actions, police departments are significantly more likely to terminate female trainees than male trainees. This predisposition toward disparate treatment raises serious questions about the even-handedness female police officer's can expect in this male stereotyped job.

Disquieting as these findings are, a note of caution in generalizing these findings is warranted. It is worth remembering that, though the facts depicted in the scenario are "real", respondents were nevertheless evaluating "paper people" in a research context. We, of course, cannot be sure that the same gender-based treatment differential occurs in the day to day operations of police departments throughout the U.S.

Given the inconsistent pattern of findings in the literature, the demonstration of gender bias in the present study suggests three potential moderators worthy of further investigation. Gender bias may be more likely a) in situations involving <u>strongly</u> gender stereotyped performance settings (e.g., dangerous situations involving physically demanding actions), b) where team performance forces the rater to apportion responsibility for outcomes among team members, and c) where performance judgments are directly linked to specific administrative actions.

WILL—Printing paid from state funds. KRB 87 378